VISIT…

Friction, Wear, and Lubrication Theory

Copyright © 1983 CRC Press LLC

THE SHAPE OF SURFACES

J. B. P. Williamson

INTRODUCTION

All surfaces are rough. The world of the engineer is made of solids whose surfaces acquire their texture as the result of a great variety of processes. In some cases it is merely a by-product of forming the bulk shape, for example, in casting, molding, or cutting. More often a separate process affecting only the surface layers is applied after the part has been formed to its bulk dimensions. Some treatments remove material, as in grinding and etching. Others, such as plating, flame spraying, and sputtering, add it. Yet others merely redistribute the surface layer: peening and calendering are examples. In addition, surfaces often show the marks of unplanned treatments such as wear and corrosion.

Surface textures found in modern engineering vary widely. Figure 1A, for example, shows a mechanically polished surface, while Figure 1B shows one which has been electroplated. Such surfaces may feel smooth and give a mirror-like reflection, yet the electronmicrographs show they are covered with hills and valleys. Figure 2 places this roughness in perspective against other surface-related phenomena of interest in engineering.

Whenever two solids are brought together, they touch first where hills on one contact the surface of the other. As the hills flatten, contact areas grow and the pressure falls until it becomes too low to cause further deformation. Contact is thus limited to a relatively small area, and the rest of the surfaces are held apart. The interfacial gap formed is usually continuous, permitting gaseous and liquid access to the whole interface (Figure 3 illustrates this). Two copper surfaces were pressed together and sulfur dioxide gas was allowed to diffuse into the interfacial gap. On separation, bright areas of intimate contact (where the copper was protected from the gas) were in clear contrast to the chemically discolored surface. Areas of contact about 1 to 5 μm across and about 10 to 50 μm apart are typical of many tribological interfaces.

The texture of a surface ranges from large-scale shape deviations to tiny features such as ledges in crystal faces and steps where dislocations emerge. The scale of the world of the tribologist is essentially determined by the size of the individual contact areas between surfaces. Features which are small compared with individual contact regions are not usually significant.

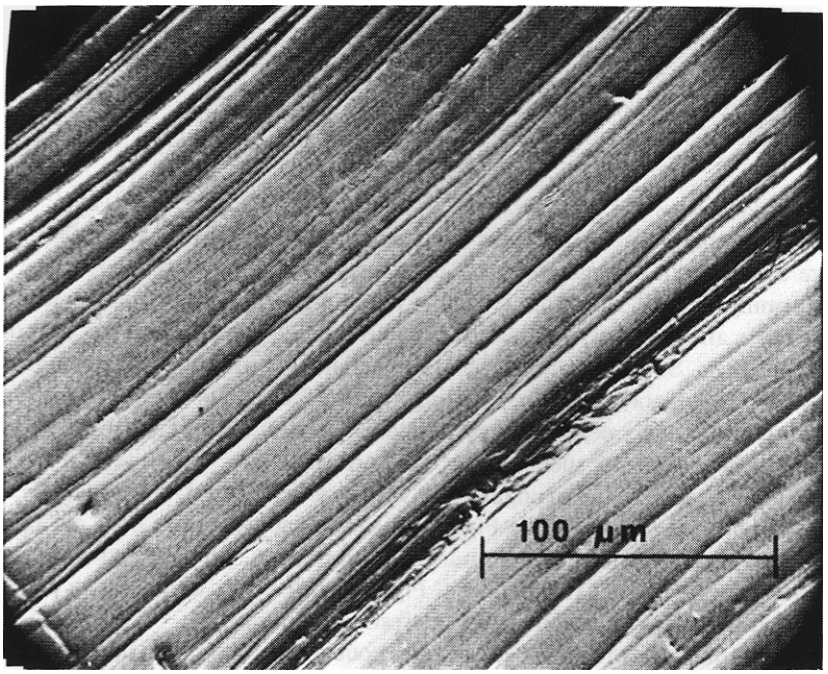
MEASUREMENT OF SURFACE ROUGHNESS

The principal instruments used to study surface shape are the scanning electron microscope (SEM) and the profile analyzer. The SEM can provide micrographs with sufficient resolution to reveal individual details and, yet, has a large enough field of view that the interrelation of many such features can be seen. In practical tribology, however, it has two disadvantages: specimen size is limited and it cannot quantify roughness.

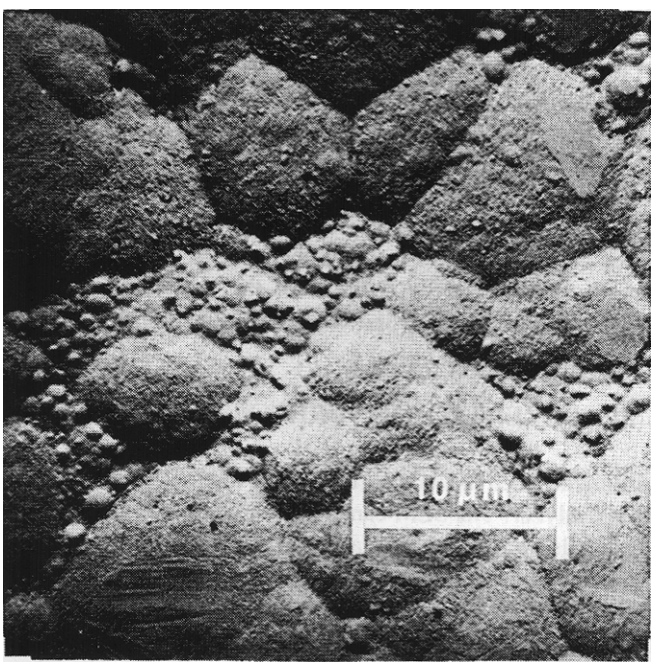
Checking a surface against a specification or measuring how texture influences performance requires numerical descriptions. The profile analyzer is the most widely used instrument for this. It draws a sharp stylus lightly over the specimen and detects its movement as it follows the texture. The signal is amplified and recorded on a chart to produce a profile of the surface. Many surfaces contain flaws — unintentional, infrequent defects, such as cracks, inclusions, and scratches. Profiles should be positioned to avoid these aberrations whenever possible.

Surface profiles usually contain three major components (Figure 4):

Copyright © 1983 CRC Press LLC



A



B

FIGURE 1. Electron micrographs of (A) mechanically polished copper surface and (B) gold electroplate deposited on brass.

Copyright © 1983 CRC Press LLC

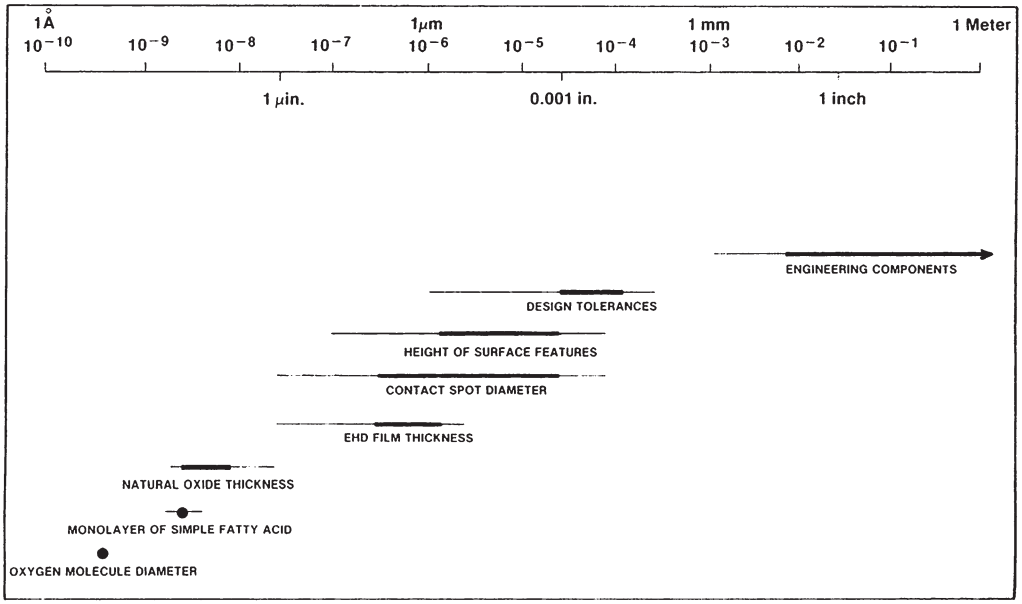


FIGURE 2. Comparative size of surface-related phenomena.

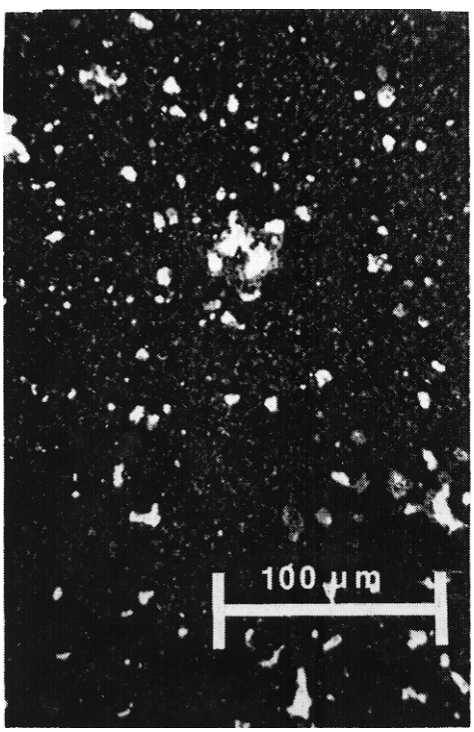


FIGURE 3. Contact areas revealed by chemical decoration. Two copper surfaces were pressed together for 1 month in an atmosphere of 1% SO_2 in air at 85% R.H., then photographed on separation.

Copyright © 1983 CRC Press LLC

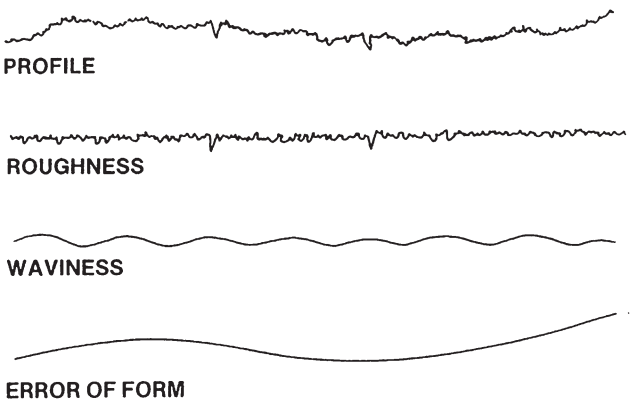


FIGURE 4. The three components of surface texture.

1. **Roughness** — closely spaced irregularities, the height, width, and direction of which create the predominant surface pattern. Roughness includes those surface features intrinsic to the production process.
2. **Waviness** — surface irregularities of greater spacing than roughness. They are often the result of heat treatment, machine or workpiece deflections, vibrations, or warping strains.
3. **Errors of form** — gross deviations from nominal shape. They are not normally considered part of the surface texture.

It is important to analyze the profile of surface features over an appropriate length. If the function of a particular surface depends primarily on its short wavelength irregularities, for instance, a misleadingly large value of roughness would be obtained if the entire profile, including the waviness, were analyzed. The greatest spacing of surface irregularities to be included is called the "roughness-width cutoff". In most modern instruments an electronic filter removes unwanted components of the profile as illustrated in Figure 5. The top profile represents the actual movement of the stylus on a surface. The lower ones show the same profile using cutoffs of 0.8, 0.25, and 0.08 mm. The international standard roughness-width cutoffs are

Millimeters	Inches
0.08	0.003
0.25	0.010
0.80	0.030

The preferred value of 0.80 mm is assumed unless a different cutoff is specified. Roughness-width cutoff is sometimes called "roughness sampling length". The distinction between the terms is a fine technical one concerning instrument design; for practical purposes they are essentially the same.

Roughness of the filtered profile is most frequently described in terms of the arithmetic average deviation from its mean, which is called R_a. It was previously called "arithmetic average" (AA) or "center line average" (CLA). Usually the profile analyzer presents R_a as the mean of several readings taken consecutively along the profile. R_a is almost universally used on engineering drawings. Table 1 gives internationally adopted values together with the alternative "roughness grade number".

A second important measure of surface texture advocated by the International Organization

Copyright © 1983 CRC Press LLC

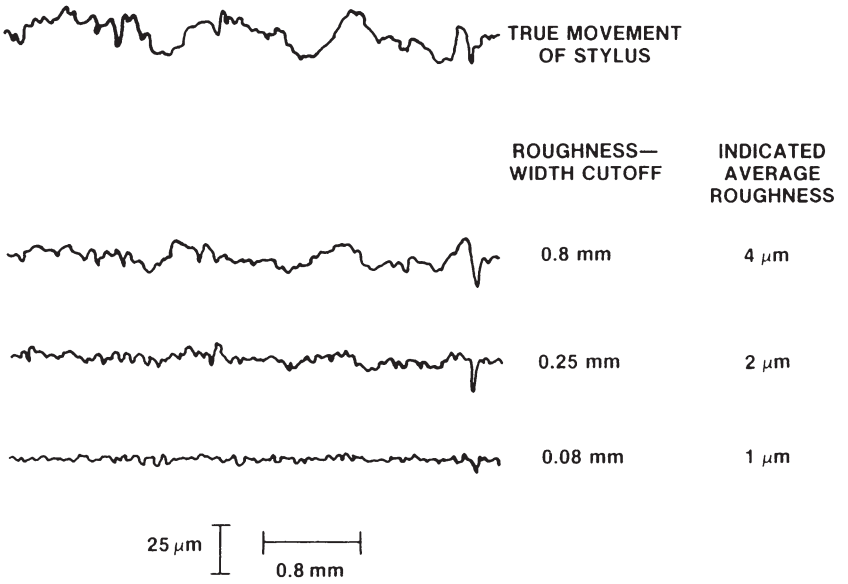


FIGURE 5. The effect of the roughness-width cutoff is to remove all components of the total profile which have wavelengths greater than the cutoff.

Table 1
ARITHMETIC AVERAGE ROUGHNESS GRADES

Recommended R_a values		Roughness grade number
μm	μ in.	
0.025	1	N1
0.05	2	N2
0.1	4	N3
0.2	8	N4
0.4	16	N5
0.8	32	N6
1.6	63	N7
3.2	125	N8
6.3	250	N9
12.5	500	N10
25	1000	N11

for Standardization (ISO) is the average elevation of the five highest peaks above the five lowest valleys, which is called R_z or the ISO 10-point height. It is measured over a single sampling length (roughness-width cutoff), and is particularly useful when only a small surface is available for assessment. R_a averages data from the entire height range of the profile, and is thus insensitive to occasional peaks or valleys. R_z, on the other hand, describes roughness in terms of extreme height and is valuable when performance of the surface could be impaired by excrescences or cracks.

R_a and R_z are easily determined and are widely used to monitor production consistency. They have one major disadvantage: both give equal weight to the shape of valleys and peaks. In practice, however, behavior of a surface in contact with another depends essentially on

Copyright © 1983 CRC Press LLC

the texture of its highest strata, and hardly at all on the shape of its valleys. Frequently the highest parts of engineering surfaces differ significantly from the general texture.

Likely severity of wear between sliding surfaces is given by the Plasticity Index, which indicates whether deformation in the contact regions will be predominantly elastic or plastic. This index is given by $(E'/H)\sqrt{(\sigma/\beta)}$, where $1/E' = (1 - \nu_1^2)/E_1 + (1 - \nu_2^2)/E_2$ and E_1,E_2, ν_1, and ν_2 are Young's moduli and Poisson's ratios of the contacting solids, H is the hardness, σ is the standard deviation of the height of the hills, and β is the mean radius of their summits. A Plasticity Index of 1 or less indicates essentially elastic contact with a low probability of wear. Values above 3 indicate mainly plastic contact regions with a higher probability of wear. Ball and roller bearings and well run-in surfaces have indices around 1; almost all other engineering surfaces have 10 or more.

Some modern surface analyzers provide a voltage analog which permits a detailed computer analysis of the surface profile. In particular, height distribution of hills and the mean radius of their summits, which are components of the Plasticity Index, can be computed. Certain roughness parameters depend on the interval at which the analog signal is digitized. The number of peaks apparent in the profile, for example, and consequently the average radius of their summits, can vary widely when this interval is changed.

Departure of the surface profile from its mean may also be expressed in terms of the root mean square average deviation. This measure of roughness, called R_q, is similar to the arithmetic average R_a, although often 10 to 20% higher. It is particularly important in the theory of surface contact. In practical engineering it is frequently, through incorrectly, used interchangeably with R_a. Skewness is a useful measure of the asymmetry of the profile. A surface which is a plateau with occasional deep rifts, a "scratchy" surface, is said to have negative skewness. A plain with ridges, a "peaky" surface, has positive skewness.

Many production processes impart a directionality, or lay, to the surface. In principle, so long as the stylus of the profile analyzer traverses a representative sample of the texture, the R_a, measured will be independent of the orientation of the track. However, the apparent wavelength of the surface features depends on the angle between the track and the lay. The orientation thus determines which features will be cut off by the filter. It is standard practice, therefore, to measure roughness at right angles to the lay, and this is assumed in specifications unless otherwise indicated.

The basic symbol used to designate surface texture is the checkmark.

√

A triangle is used when the surface is to be machined.

∇

A circle means that the texture must be produced without any bulk removal of material.

Various numbers and symbols may be written against the checkmark to specify features of the texture, the most common are given below.

The maximum R_a acceptable is shown in micrometers.

Copyright © 1983 CRC Press LLC

If necessary, the minimum acceptable R_a is added.

A horizontal line is added to the checkmark to specify further features as in the examples given below.

The maximum acceptable waviness height and spacing (peak to valley height and peak to peak distance of the waviness) are shown in millimeters. The height is written first.

The roughness-width cutoff to be used in the measurement is shown in millimeters. When no value is given 0.8 is assumed.

The lay required is indicated by a lay symbol placed thus:

Standard lay symbols are

Symbol	Interpretation
=	Approximately parallel to the line representing the surface to which the symbol is applied
⊥	Approximately perpendicular to the line representing the surface to which the symbol is applied
X	Angular in both directions to the line representing the surface to which the symbol is applied
M	Multidirectional
C	Approximately circular relative to the center of the surface to which the symbol is applied
R	Approximately radial relative to the center of the surface to which the symbol is applied
P	No lay, e.g., pitted, protuberant, particulate, or porous

CARTOGRAPHY OF SURFACE ROUGHNESS

The shape of a surface may be displayed by a computer-generated map developed from digital data derived from many closely spaced parallel profiles. Such a map shows details of individual features and also the general topography over a relatively large area. While these maps are tedious to obtain, the advent of computer-coupled profile analyzers will encourage wider use of this potent method of describing surfaces. Figure 6 shows part of such a map of a bead-blasted surface.

There is often remarkable similarity between maps of the surface of solids and ordinary contour maps of the surface of the earth. The scale factor is about 10^8 to 1. The ratio of height to spacing of the hills are similar. The slopes, in both cases, are usually between 1 and 10° and are rarely steeper than 30°. Take, for example, the mountains of New England as a model of a metal surface. A naturally occurring oxide layer would then be represented to scale by a 3-ft snowfall; an oxygen molecule by a golf ball; and a monolayer of a simple fatty acid, such as stearic acid, by a covering of 1-ft high grass. On the same scale an engineering component a few inches across would be the North American continent, and individual areas of contact would be a little larger than football stadiums.

If maps are made of two surfaces which are to be placed in contact, the gap between

Copyright © 1983 CRC Press LLC

FIGURE 6. A glass surface lightly bead-blasted. Map made from height data collected from many parallel closely spaced profiles. The roughness is approximately 1 μm R_a. The white areas are plateau of the original glass surface. The countours are every micrometer.

Copyright © 1983 CRC Press LLC

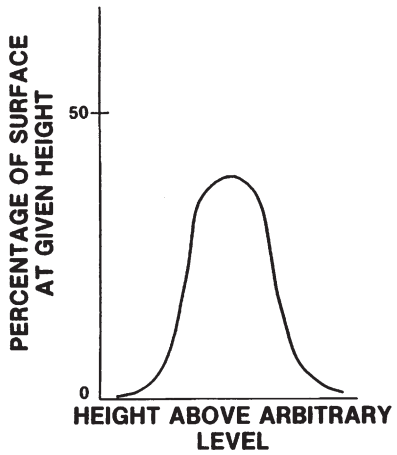


FIGURE 7. A typical height distribution (ADF) for a near-Gaussian surface.

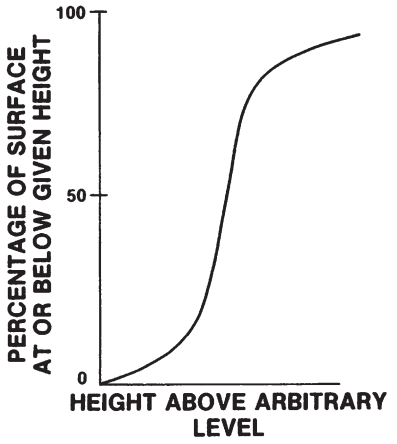


FIGURE 8. Cumulative presentation of Figure 7: the complement of the bearing area curve (BAC).

them at each point can be determined and printed by the computer. Contours of this gap-map indicate the areas of contact between the surfaces at different loads. Analyses of gap-maps suggest normal contact is almost entirely confined to the highest 25% of each surface, and mainly occurs in the top 10%. (The percentages are of surface area, not of height.)

HEIGHT DISTRIBUTION AND BEARING AREA CURVES (BAC)

The fraction of a surface lying in each stratum can readily be obtained by tracing a profile and sampling its height at regular intervals. This gives the height distribution, sometimes called the "amplitude density function" (ADF). The profile must be long compared with the surface irregularities and include a representative sample of the texture. Figure 7 shows a typical height distribution.

It is, however, more useful to describe surfaces in terms of the integral of the distribution (Figure 8), which gives the fraction of the surface at or below each height. The well-known complementary BAC, which gives the area of contact which would exist if the hills were worn down to the given height by an ideally flat body, is the fraction of the surface at or above each height. Some modern surface analyzers provide chart or video displays of height distributions and BAC as standard features.

Gaussian Height Distribution

Many engineering surfaces have height distributions which are approximately Gaussian, i.e., they can be described by the normal probability function. A Gaussian distribution plotted on probability graph paper appears as a straight line, and deviations are readily detected. A sample of several thousand height readings will normally be needed, requiring data from many profiles (which must share a common reference level).

Figure 9 gives the BAC of a bead-blasted aluminum surface. More than 22,000 height readings show that at most only the top 0.01% of the surface was non-Gaussian. The slope of the line is a direct measure of the roughness parameter R_q. To an adequate approximation, R_q is 0.6 times the height difference between the 20th and 80th percentiles of the surface; if a longer straight line is available, R_q can be taken as 0.3 times the height difference between the 5th and 95th percentiles. For a surface with a Gaussian height distribution, R_a is approximately 0.8 R_q. In Figure 9, R_q is 1.25 μm and R_a is 1.0 μm. In general, however, there is no simple relation between R_a and R_q.

Copyright © 1983 CRC Press LLC

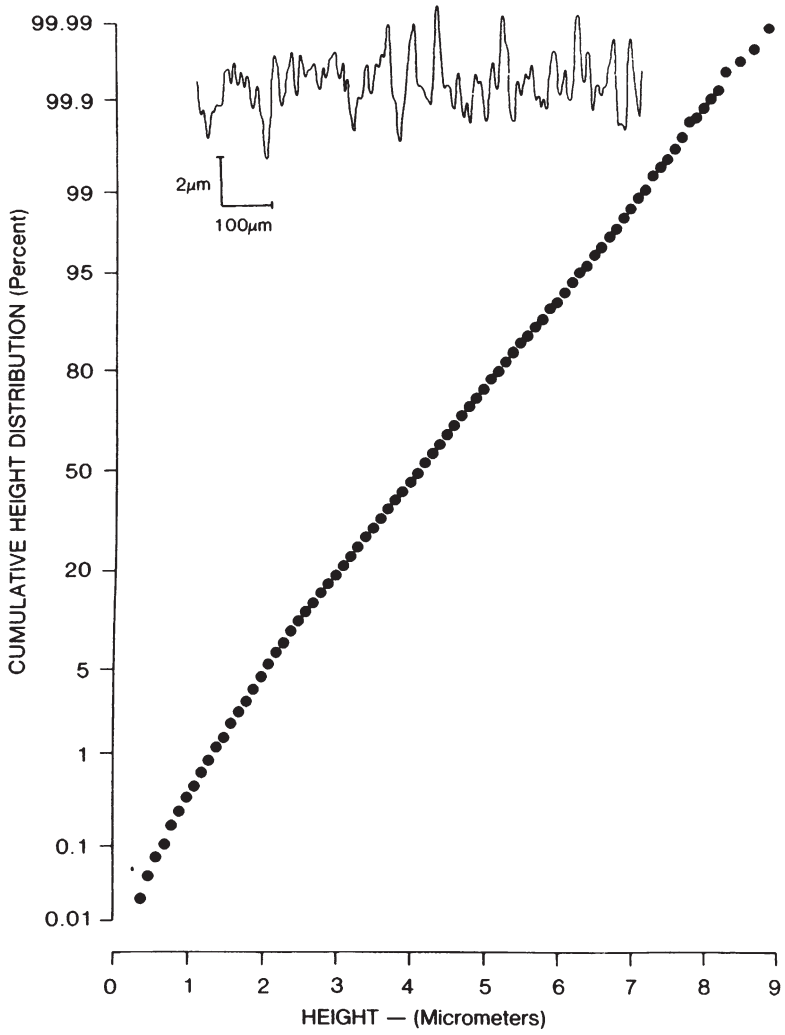


FIGURE 9. Cumulative height distribution of a bead-blasted surface. The surface distribution is Gaussian — at the high end to at least 3.5 standard deviations from the mean. A profile of the same surface is also shown: its vertical magnification is 50 times the horizontal.

Several finishing processes create Gaussian height distributions. In these treatments the texture results from a very large number of separate events each of which affects only a small region of the surface, e.g., the impact of one bead during blasting, the deposition of one ion in plating, or a single contact with a particle of polishing powder. If such a preparation technique involves a large number of randomly located events, the Central Limit Theorem of statistics says that the shape will tend towards a Gaussian height distribution. It is not even necessary that the events are all the same — only that their effect is cumulative and that they are numerous and randomly located. Any process satisfying these requirements will create Gaussian surfaces, regardless of the properties of the solid.

Non-Gaussian Height Distributions

In some surface treatments the formative events, though numerous and cumulative, do not occur randomly. Their location may depend on the microstructure of the solid, as in

Copyright © 1983 CRC Press LLC

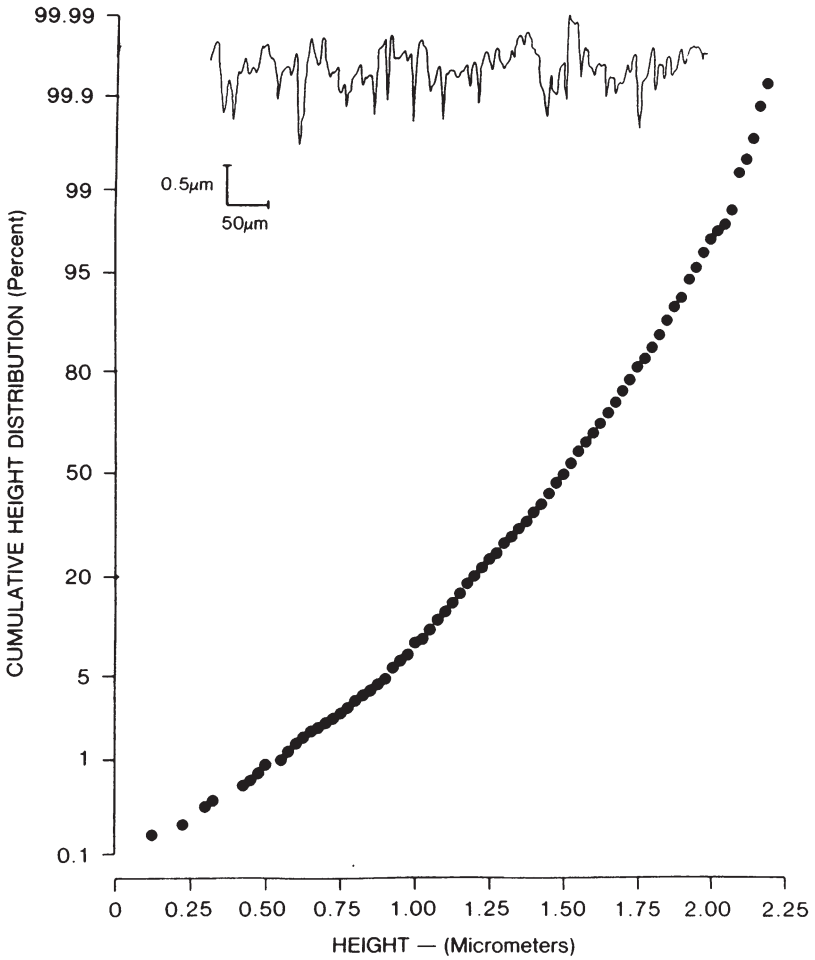


FIGURE 10. Cumulative height distribution and typical profile of a ground surface. The distribution appears on probability paper as a gentle curve. This is characteristic of surfaces formed by extreme event processes.

etching, in which material removal is influenced by alloy phases, grain orientation, and grain boundaries. Alternatively, a process may be nonrandom because it is influenced by existing topography. Electroplating, for example, may deposit preferentially on hills.

Other surface treatments are not cumulative. In "extreme-value" processes each region of the final surface reflects only the most extreme events which occurred there. Grinding is an extreme-value process (Figure 10). When the individual events are not numerous (during turning each point on the specimen experiences only one formative event), there is again no statistical reason for a Gaussian surface. Height distribution of a turned surface is far from Gaussian (Figure 11). The distribution of *peak* heights, however, is often close to Gaussian. This reflects the randomness of the tearing where the edge of the tool cuts the wall left by the previous pass.

PURE, MIXED, AND STRATIFIED SURFACE TEXTURES

When all features on a surface result from the same treatment, the texture is said to be pure. Such textures are created only by processes which obliterate all previous treatments (milling or melting, or roughening a much smoother surface, for example).

Copyright © 1983 CRC Press LLC

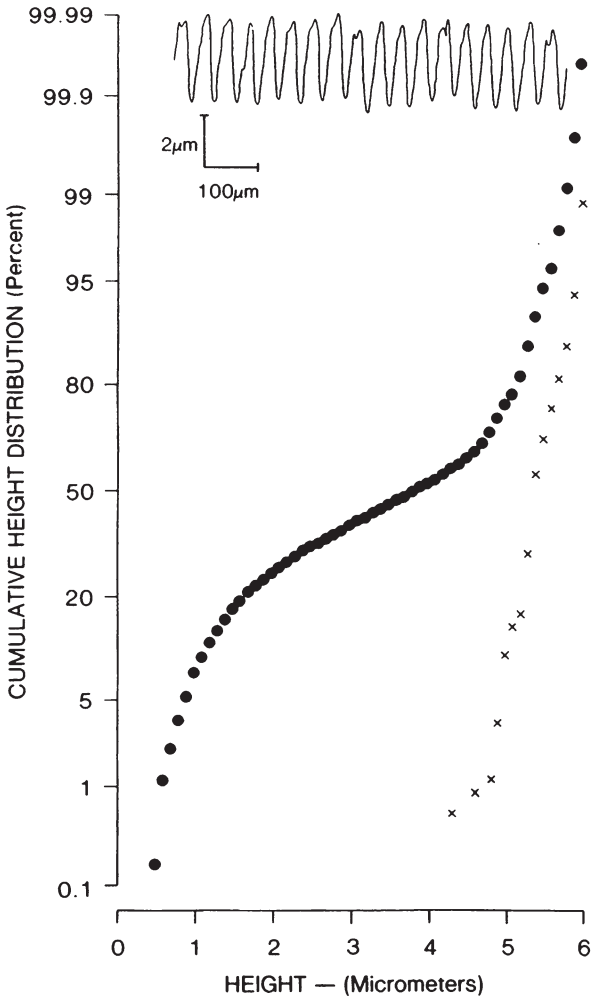


FIGURE 11. Cumulative height distribution and typical profile of a turned surface (solid circles). The distribution is definitely non-Gaussian. The profile peaks (x's) however are much more nearly Gaussian.

Most surfaces have mixed textures as evidence of more than one preparation process. Normally features generated by the second process are distributed randomly, so that every region of the finished surface bears the marks of both treatments. However, some treatments give rise to "stratified surfaces" in which the valleys reflect one preparation process while the hills have a completely different texture from another treatment. R_a and R_z can then be very misleading; height distribution plotted on probability paper is much more useful.

Surfaces produced by running-in usually have stratified textures. Figure 12 follows such a process in which a mild steel pad lubricated with SAE 20 oil was worn against a finely ground hard steel flat. The solid circles show that before wearing the pad had a Gaussian height distribution with an R_a of 1.25 μm (given by the slope of the line). The open circles show how the texture changed as the surface wore. A typical run-in surface has a stratified texture such as that shown by curve B. The upper texture (which determines the nature of the contact) will behave as though it were part of a smooth surface. The interfacial gap (which controls lubricant access and removal of debris) depends on the lower texture.

Copyright © 1983 CRC Press LLC

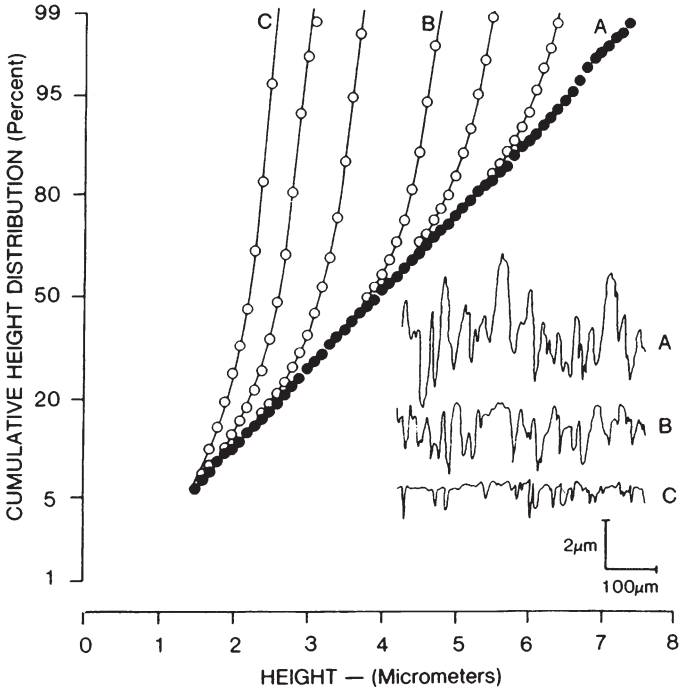


FIGURE 12. The effect of wear. The initial height distribution (A) and six non-Gaussian distributions (open circles) of a bead-blasted surface represent, from right to left, progressive slages. Height distributions of this form are typical of those created by stratified secondary preparation processes.

Surfaces with mixed textures can be designed for particular combinations of properties which cannot be achieved with a single surface treatment. The surfaces of heavy-duty sliding bearings, for example, can be made to have wide interfacial gaps to facilitate lubricant access and debris removal; with a second process, they can also be given shallow-domed plateaus which can carry the load with very little plastic deformation, and hence little wear. Conversely, seal surfaces can be honed or lapped and then given noninterconnecting dimples to carry some lubricant to reduce wear.

REFERENCES

1. American National Standard ANSI B46.1, Surface Texture, 1978.
2. American National Standard ANSI Y14.36, Surface Texture Symbols, 1978.
3. British Standard BS 1134, Assessment of Surface Texture, 1972.
4. Canadian Standard CSA B95, 1962.
5. International Standard ISO R468, Surface Roughness, 1974.
6. **Archard, J, F.,** Surface topography and tribology. *Tribol. Int.,* 213-221, 1974.
7. **McAdams, H. T.,** Quantitative microcartography: a physico-geometric approach to surface integrity, in *Proc. Int. Conf. Surface Technology, Pittsburgh,* Society of Manufacturing Engineers, Dearborn, Mich., 1973, 96.
8. **Sayles, R. S. and Thomas, T. R.,** Mapping a small area of a surface, *J. Phys. E.,* 9, 855, 1976.
9. **Tabor, D.,** The solid surface, *Phys. Bull.,* 29, 521, 1978.

Copyright © 1983 CRC Press LLC

10. **Whitehouse, D. J.,** Surfaces — a link between manufacture and function, *Proc. Inst. Mech. Eng. (London),* 192, 179, 1978.
11. **Whitehouse, D. J.,** Stylus techniques, in *Characterization of Solid Surfaces,* Kane, P. F. and Larrabee, G. R., Plenum Press, New York, 1974, chap. 3.
12. **Williamson, J. B. P.,** Microtopography of surfaces, *Proc. Inst. Mech. Eng. (London),* 182,(Pt. 3K), 21, 1967.
13. **Williamson, J. B. P., Pullen, J., and Hunt, R. T.,** The shape of solid surfaces, in *Surface Mechanics,* American Society of Mechanical Engineers, New York, 1969, 24.
14. **Young, R. D.,** Surface microtopography, *Phys. Today,* 42-49, November 1971.

Copyright © 1983 CRC Press LLC